Impact of Perturbations on Nonlinear Frequency-Division Multiplexing

Xianhe Yangzhang, Domaniç Lavery, *Member, IEEE*, Polina Bayvel, *Fellow, IEEE*, and Mansoor I. Yousefi

(Invited Paper)

Abstract—Nonlinear frequency-division multiplexing (NFDM) is a communication scheme in which users' signals are multiplexed in the nonlinear Fourier domain. The contributions of this paper are twofold. First, the achievable information rates (AIRs) of NFDM based on an integrable model of the optical fiber are summarized. For this ideal model, it is shown that the AIR of the NFDM is greater than the AIR of the wavelength-division multiplexing (WDM) for a given bandwidth and signal power, in a representative system with five users and one symbol per user. The improvement results from nonlinear signal multiplexing. Second, the impact of some of the main perturbations on NFDM are investigated, including the fiber loss, polarization effects and the third-order dispersion. For a realistic non-ideal model, it is shown that the WDM AIR with joint dual-polarization back-propagation and third-order dispersion compensation is approximately equal to the NFDM AIR with two independent single-polarization demodulations and without third-order dispersion compensation. Using a joint dual-polarization receiver and perturbations compensation is expected to increase the NFDM AIR.

Index Terms—Nonlinear Fourier transform, nonlinear frequency-division multiplexing, wavelength-division multiplexing, optical fiber communication.

I. INTRODUCTION

OPTICAL communication systems have experienced an extraordinary increase in data rates in the past few decades. One of the enabling technologies is wavelength-division multiplexing (WDM), which makes efficient use of the available fiber bandwidth. However, driven by the increased traffic demand, WDM systems are now approaching their theoretical limits set by fiber nonlinearity.

It is well known that nonlinear interactions pose a limitation to the achievable information rates (AIRs) in WDM optical networks [1]. Nonlinear frequency-division multiplexing (NFDM), a signal multiplexing scheme based on the nonlinear Fourier transform (NFT), is a promising approach to overcome the "capacity crunch" problem in WDM optical systems [2]–[4].

The nonlinear Fourier spectrum consists of a continuous component and a discrete component. Information transmission using the continuous spectrum is studied in [4]–[11], while discrete spectrum modulation is studied in [4], [12]–[14]. Recent experimental demonstrations of data transmission based on the NFT include joint discrete and continuous spectrum modulation [8]. A record data rate of 32 Gb/s was demonstrated at OFC 2017 using the continuous spectrum of NFT and the 32QAM modulation format. A peak-SNR gain of 1.3 dB was achieved over a comparable OFDM system [15]; for comparison, see also [16]–[22].

Previous works have mostly applied the NFT to point-to-point links, often as a nonlinear compensation scheme. However, the main potential of NFDM is realized in network environments. Here, users' signals are multiplexed in the nonlinear Fourier spectrum in disjoint intervals and, in the absence of noise, propagate independently in the network. Crucially, the signals of the user-of-interest (UOI) will not be distorted by co-propagating signals. As a result, the deterministic inter-symbol and inter-user interference are simultaneously zero for all users of a network.

Research in NFDM initially demonstrated proof-of-concepts, showing how this scheme works in point-to-point channels. However, the AIRs of the NFDM signals were at best the same as the AIRs of the WDM signals in these initial demonstrations [4]. Advances in numerical methods made it possible to multiplex signals in the nonlinear Fourier domain and explore the NFT at high powers. The AIRs of NFDM and WDM were compared for the first time recently in [23], [24] for fiber in the defocusing regime. It was shown that NFDM achieves data rates higher than WDM rates, subject to the same power and bandwidth constraints, in illustrative systems.

The objective of this paper is twofold. First, we summarize the AIRs of NFDM, compared to WDM, for a given input power and bandwidth in an ideal integrable model in the focusing regime. For this model, it is shown that the NFDM AIR is greater than the WDM AIR in a representative system with five users and one symbol per use. This part summarizes the recent conference paper on NFDM [25]. One of the main limitations of NFDM is that it is based on ideal integrable models of the optical fiber, for example, with ideal distributed Raman amplification. It is not clear how NFDM performs in realistic systems with non-idealities and perturbations. Second, we study the impact of the following perturbations on NFDM: fiber loss, periodic amplification, third-order dispersion and polarization effects. It is shown that using a path-averaged loss model, the impact of the attenuation (with periodic EDFA amplification) on NFDM is small. It is shown that the WDM AIR with joint dual-polarization back-propagation and perturbations compensation is nearly equal to the NFDM AIR with two independent single-polarization demodulations

Xianhe Yangzhang, Domaniç Lavery and Polina Bayvel are with the Department of Electrical and Electronic Engineering, University College London, WC1E 7JE London, UK (e-mail: x.yangzhang@ucl.ac.uk)

Mansoor I. Yousefi is with Télécom Paris Tech, 75013 Paris, France (email: yousefi@telecom-paristech.fr).

and without perturbations compensation, in a representative system with five users and one symbol per user. With joint dual-polarization transmission and compensation of the same perturbations in both schemes, the NFDM AIR is expected to be higher in the considered system. We study the dependency of the AIRs with the number of symbols. As the number of symbols is increased, the AIRs of both WDM and NFDM are decreased. The results of the paper clarify to what extent NFDM applies to realistic systems.

The paper is presented as follows. In Sec. II, we review the channel model and explain the origin of the data rate limitations in today's optical fiber networks. The NFDM theory is shortly summarized in Sec. III. The achievable rates of NFDM and WDM are computed and compared in Sec. IV, in a multi-user system with five users and one symbol per user. In Sec. V, NFDM is studied for non-integrable models with non-idealities and perturbations. After comments on the complexity of the NFT in Sec. VI, the paper is concluded in Sec. VII.

II. FIBER-OPTIC NETWORKS

A. Channel Model

Signal propagation in single-mode single-polarization optical fiber with ideal distributed Raman amplification can be modeled by the stochastic nonlinear Schrödinger equation. The equation in the normalized form reads

$$j\frac{\partial q}{\partial z} = \frac{\partial^2 q}{\partial t^2} - 2s|q|^2 q + n(t, z), \quad (1)$$

where $j = \sqrt{-1}$ and $q(t, z)$ is the complex envelope of the signal as a function of time t and distance z along the fiber, $n(t, z)$ is white Gaussian noise with power spectral density σ_0^2, and $s = 1$ in the defocusing regime (normal dispersion fiber), while $s = -1$ in the focusing regime (anomalous dispersion fiber). The signal and noise are band-limited to the same bandwidth B for all $0 \leq z \leq 1$. The reader is referred to [2] for further details about the model, such as the normalization procedure.

In this paper, we consider a *network environment*. This is an optical network with the following set of assumptions: (a) there are multiple transceiver pairs, (b) there are add-drop multiplexers (ADMs) in the network; the locations and the number of ADMs are unknown, (c) each transceiver pair does not have information of the incoming and outgoing signals in the path that connects them. In a network environment, the signal of the user-of-interest co-propagates with the signals of the other transceiver pairs in the network and, in the case of WDM, is subject to inter-user interference [4, Sec. II. B.3.]. A typical network environment is depicted in Fig. 1.

B. Capacity Limits of the WDM Networks

WDM systems modulate information in the Fourier basis corresponding to different frequencies or wavelengths. The basis elements could be modulated independently at the transmitter, but they couple together in the nonlinear fiber channel. Consequently, the optical fiber channel can be viewed as a multi-user interference channel (IC). The capacity of the
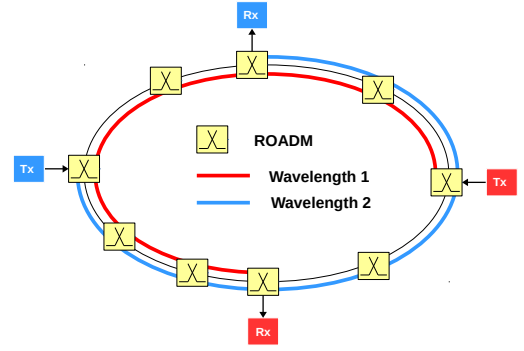

Fig. 1. A network environment with two transceivers.

optical fiber channel is therefore not captured by a single number. However, in practice, a useful quantity is the sum rate. Several distinct strategies exist for communication over an IC: (i) treating interference as noise, (ii) orthogonalization and (iii) interference alignment. An extensive body of work exists presenting the AIRs of WDM optical networks, which are sometimes referred to as the "nonlinear Shannon limit" [1], [26]–[29]. However, these rates are lower bounds on the capacity of the user-of-interest, implicitly using strategy (i). Other strategies for WDM channels, such as (ii) and (iii), exist as well, which predict different AIRs.

NFDM is an approach based on strategy (ii), i.e., while WDM treats interference as noise, NFDM is an approach to channel orthogonalization. We will show in Sec. IV that the AIR of strategy (i) (WDM) is strictly lower than the AIR of strategy (ii) (NFDM) at high powers.

III. REVIEW OF NFDM

We briefly review basic NFDM theory from [23]. The theory in this section applies to both focusing and defocusing regimes. However, the simulations in the subsequent sections are presented only for the focusing regime.

Let $\hat{q}(\lambda, z) : \mathbb{R} \times \mathbb{R}^+ \mapsto \mathbb{D}$ be the nonlinear Fourier transform of $q(t, z)$ as a function of the nonlinear frequency λ and distance z, where, in the focusing regime, $\mathbb{D}$ is the complex plane $\mathbb{C}$ and, in the defocusing regime, is the unit disk $\mathbb{T} = \{z \in \mathbb{C} : |z| < 1\}$.

In the defocusing regime $|\hat{q}(\lambda, z)| < 1$. We thus introduce the following transformation to map functions with co-domain $\mathbb{D}$ to functions with co-domain $\mathbb{C}$:

$$U(\lambda, z) = \left(-s \log(1 - s|\hat{q}(\lambda, z)|^2) \right)^{\frac{1}{2}} e^{j \angle \hat{q}(\lambda, z)}, \quad (2)$$

where $\angle \hat{q}$ is the phase of $\hat{q}$. This eliminates the unit peak power constraint in the defocusing regime. We use the transformation (2) in both focusing and defocusing regimes. In the defocusing regime, it is necessary so that the signal space is a vector space. In the focusing regime it is not necessary, but it makes signal energies in the time and nonlinear Fourier domain the same, which is useful.
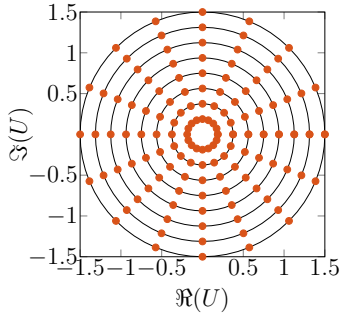
Fig. 2. Uniformly-spaced multi-ring constellation in the U-domain

We define a *generalized time* τ that is related via the Fourier transform to the nonlinear frequency $\lambda/2\pi$. Let $u(\tau, z) \leftrightarrow U(\lambda, z)$ be a Fourier transform pair, i.e.,

$$u(\tau, z) = \sqrt{2}\mathcal{F}^{-1}\{U(\lambda, z)\}. \qquad (3)$$

The factor $\sqrt{2}$ is introduced so that the energy (computed according to the formulas in [23]) of $u(\tau, z)$, $U(\lambda, z)$, $\hat{q}(\lambda, z)$ and $q(t, z)$ are the same; see [23, Eq. 28].

A. NFDM Transmitter

Consider a multi-user system with N_u users, each having N_s symbols, with total linear bandwidth B and total average power $\mathcal{P}$. Modulation in NFDM begins by forming

$$u(\tau, 0) = \sqrt{2} \sum_{k=-\frac{N_u}{2}}^{\frac{N_u}{2}-1} \sum_{\ell=-\frac{N_s}{2}}^{\frac{N_s}{2}-1} s_\ell^k \phi(\tau - \ell T_0)e^{j2\pi k W_0\tau}, \qquad (4)$$

where $\phi(\tau)$ is a root-raised-cosine function with unit energy, bandwidth W_0 and the roll-off factor r, $T_0 = 1/W_0$, $\{s_\ell^k\}_\ell$ are the symbols of user k at time instance ℓ chosen from a multi-ring constellation Ξ; see Fig. 2.

Next $U(\lambda, 0)$ is computed according to (3) and subsequently $\hat{q}(\lambda, 0)$ is obtained as

$$\hat{q}(\lambda, 0) = \left(s - se^{-s|U(\lambda,0)|^2}\right)^{\frac{1}{2}}e^{j\angle U(\lambda,0)}.$$

Finally, we have

$$q(t, 0) = \text{INFT}\{\hat{q}(\lambda, 0)\}.$$

B. NFDM Receiver

At the receiver, forward NFT is applied to obtain $\hat{q}(\lambda, \mathcal{L}) = \text{NFT}\{q(t, \mathcal{L})\}$. Then $U(\lambda, \mathcal{L})$ and $u(\tau, \mathcal{L})$ are computed using (2) and (3). The received symbols are then obtained by matched filtering

$$\hat{s}_\ell^k = \int_\infty^{-\infty} u(\tau, \mathcal{L})\phi^*(\tau - \ell T_0)e^{-j2\pi k W_0\tau} d\tau.$$

C. Channel Filter

An important property of the NFT, crucial to the communication problem, is that, in the absence of noise, $\hat{q}(\lambda, z)$ propagates in distance according to an all-pass-like filter

$$\hat{q}(\lambda, z) = e^{4js\lambda^2 z}\hat{q}(\lambda, 0). \qquad (5)$$

It can be seen that signal propagation in the nonlinear Fourier domain is governed by simple multiplication of a filter. Importantly, the propagation of different nonlinear frequency components is independent of one another, which is the reason that multiplexing in the nonlinear Fourier domain is of interest. The input-output relation (5) is similar to the way that signals in the (linear) frequency domain propagate in the orthogonal frequency-division multiplexing (OFDM).

IV. COMPARISON OF THE ACHIEVABLE RATES OF WDM AND NFDM

In this section, we compare the achievable rates of NFDM and WDM under the same bandwidth and power constraints in the focusing regime (corresponding to the standard single-mode fiber). The material in this section is a summary of the conference paper [25].

A. AIRs of WDM and NFDM

First a simple multi-user simulation for one signal without noise is presented to show the origin of the NFDM gain. We consider a 5-user NFDM and 5-user WDM system with the same overall linear bandwidth of 100 GHz and total signal power $P = E/T$, where E and T are the energy and time duration of the (entire) multiplexed signal $q(t, 0)$. Time duration and bandwidth are defined as intervals that contain 99% of the signal energy. Root-raised-cosine pulses with roll off factors of 25% and standard single-mode fiber are considered. We consider one symbol per user with fiber parameters shown in Tab. I.

Fig. 3(a) shows a WDM signal at the transmitter (Tx) and receiver (Rx), with total power of 8 dBm (at Tx). The WDM receiver first filters the UOI and then digital back propagation (BP) is applied to reverse the intra-channel inter-actions. The relative error in the frequency domain is defined as $e = \|\tilde{q}(f, z)) - \tilde{q}(f, 0)\|/\|\tilde{q}(f, 0)\|$, where $\tilde{q}(f, t) = \mathcal{F}(q(t, z))$ after equalization. The mismatch error for the UOI in WDM is 95%. The resulting mismatch is the component of the non-linear interactions that cannot be compensated for in network environments using linear multiplexing. In contrast, NFDM users do not interfere with one another as shown in Fig. 3(b). The NFDM signals can be recovered almost perfectly with a relative error $e = 10^{-3} - 10^{-9}$. The Fourier spectra of signals in both systems are shown in Fig. 3(c). It can be seen that the amount of the spectral broadening in WDM and NFDM is the same.

We now turn to the main simulations presenting AIRs. The mutual information between the input output matrix of symbols (using WDM or NFDM) can be lower bounded using the chain rule for mutual information

$$I(\{s_\ell^k\}; \{\hat{s}_\ell^k\}) > I(s_0^0; \hat{s}_0^0), \qquad (6)$$

where $\{s_\ell^k\}$ and $\{\hat{s}_\ell^k\}$ are defined in Sec. III. For channels with memory, evaluating the left-hand side of (6) is not computationally feasible. The right-hand side, however, can be computed, which yields a lower bound on the capacity.

The deterministic component of the channel memory (intra-channel interactions) is equalized at the receiver in both
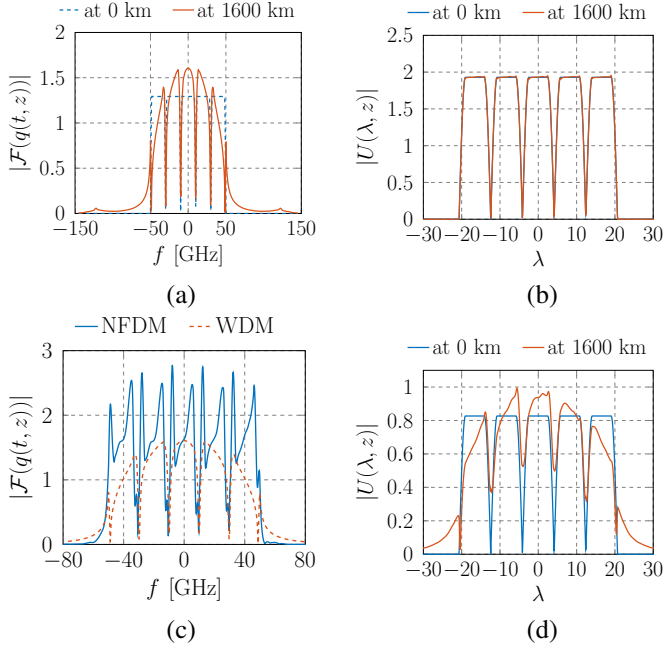
Fig. 3. For the integrable model: (a) interference in WDM; (b) lack of interference in NFDM; (c) Fourier spectra at distances with maximum input-output bandwidths. For the lossy model with periodic amplification: (d) interference in NFDM. Noise is set to zero in these figures.

TABLE I
FIBER AND SYSTEM PARAMETERS

n_{sp}	1.1	spontaneous emission factor
h	$6.626 \times 10^{-34} J \cdot s$	Planck's constant
ν	193.44 THz	center carrier frequency
α	0.046 km^{-1}	fiber loss
γ	1.27 (W.km)$^{-1}$	nonlinearity parameter
D	17 ps/(nm-km)	dispersion parameter ($s = -1$)
$\mathcal{L}$	2000 km	transmission distance
$\mathcal{W}$	100 GHz	total linear bandwidth
R_o	3	oversampling rate
$\mathcal{W}_u$	20 GHz	per-user linear bandwidth
$\mathcal{N}_u$	5	number of users
$\mathcal{N}$	2^{14}	number of samples per frame
r	0.25	roll-off factor of the raised-cosine

WDM and NFDM systems, while the stochastic component is left untreated. Computing the per-sample mutual information corresponds to ignoring the stochastic memory, which could potentially be exploited; see Sec. IV-C.

Symbols s_ℓ^k are chosen from a uniformly-spaced multi-ring constellation Ξ in the U domain (leading to a geometrically-spaced constellation in the $\hat{q}$ domain). The constellation Ξ consists of at most 64 rings and 128 phase points on each ring (13 bits per symbol). The rotational invariance of both the channel and Ξ is used to reduce the simulation time.

Noise is introduced in a distributed manner along the fiber. The noise bandwidth is set to be the maximum of the input-output signal bandwidth corresponding to the highest energy signal. The power spectral density σ_0^2 of the noise arising from distributed Raman amplification is the constant $n_{sp} h_p \nu \alpha$ for all z and in-band frequencies, with parameters in Tab. I.

The number of samples in time, linear and nonlinear spectrum is 2^{14}. The conditional probability density function (PDF)
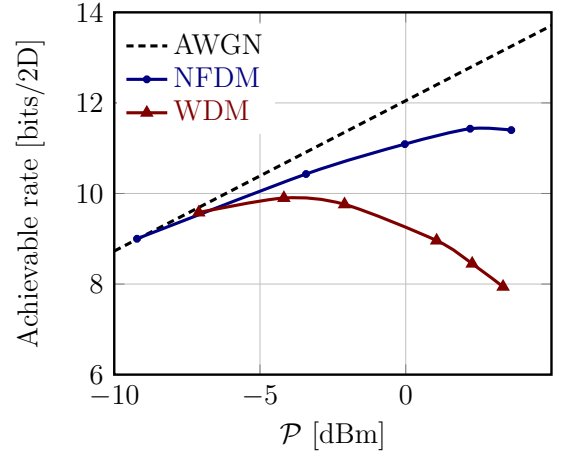

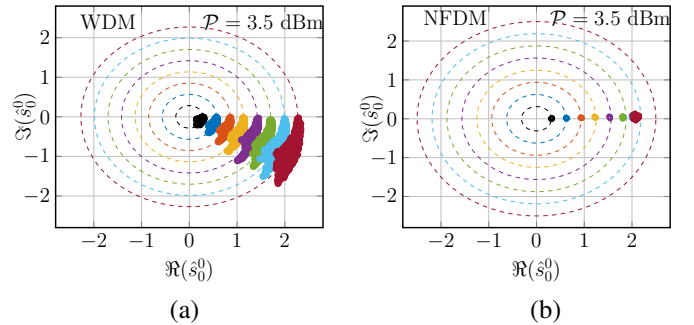Fig. 4. WDM and NFDM AIRs in the integrable model and focusing regime.

Fig. 5. Constellation in (a) WDM and (b) NFDM in the focusing regime.

$p(\hat{s}_0^0 | s_0^0)$ of the central channel $s_0^0 \mapsto \hat{s}_0^0$ is calculated based on 4500 noise realizations. The mutual information is then maximized using the Arimoto-Blahut algorithm.

The AIRs are plotted in Fig. 4 for the focusing regime. It can be seen that the AIR of WDM reaches its maximum of 10 bits/2D at a power of -4.5 dBm. In contrast, the AIR of NFDM reaches a higher maximum of around 11.5 bits/2D at 1 dBm. The apparent saturation of the NFDM AIR is attributed to neglecting the stochastic memory and numerical error in the NFT algorithm at high powers, and the potentially non-optimal constellation; see Sec. IV-C for the explanation.

The received constellations are compared in Fig. 5. It can be seen that the WDM constellation suffers from significant phase distortion. There is also a phase rotation, due to cross-phase modulation, whose average is approximately $\gamma \mathcal{L} \mathcal{P}$ [30]. On the other hand, the NFDM received constellation is similar to that in an AWGN channel [31]: the noise variances (i.e., noise 'clouds') are comparably small and approximately equal. The NFDM scheme is theoretically limited only by signal-noise interactions, which are observed to be small from Fig. 5 (b).

The conditional PDFs $p(|\hat{s}_0^0| \, \big| \, |s_0^0|)$ are plotted in [23, Fig. 9 (b)] for several values of $|s_0^0|$. The conditional PDF in NFDM is nearly independent of $|s_0^0|$, suggesting that the channel in the nonlinear Fourier domain is approximately an AWGN channel. The conditional PDF in WDM is much more diffused, leading to a higher conditional entropy and a lower AIR for WDM.

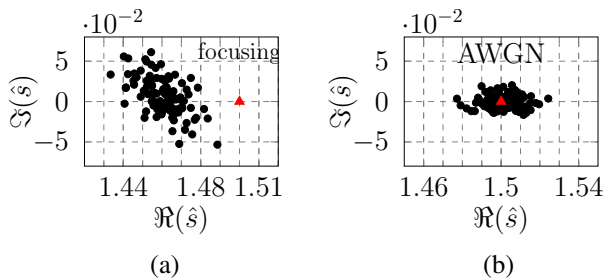The AIRs of NFDM and WDM in the defocusing regime

Fig. 6. Transmitted (red triangle) and received (black points) symbols, showing the SNR loss due to the stochastic memory.

are qualitatively similar to the AIRs in the focusing regime; the reader is referred to [24, Fig. 6].

We close this subsection by emphasizing that, the result that the NFDM-AIR is more than WDM-AIR in Fig. 4 is obtained for $N_s = 1$ and for an integrable model with single polarization. The AIRs change with N_s, as shown in [32] for the optimal scheme (see the dependency of the capacity on n in [32, Theorem 1]), and in Sec. V-C for NFDM and WDM with $N_s = 15$. As a result, the above conclusion remains to be verified for parameters not considered in this paper, particularly when $N_s > 15$.

B. Reducing the Peak-to-Average-Power-Ratio (PAPR)

The NFT signals may have high PAPR in the time domain, requiring a large number of samples to accurately represent. We use a simple pre-equalization to reduce the PAPR and the number of samples.

The idea is similar to the split digital back-propagation in [33]. Recall that the NFT spectrum evolves according to the all-pass-like filter (5). We multiply the NFT spectrum by $e^{4j\lambda^2 l}$ at the transmitter for a suitable l, to broaden the signal in the time domain. A similar idea is proposed by Tavakkolnia et al. [34]. The two approaches are related and are compared schematically in Fig. 7. In the simulations presented in this paper, we used our proposed PAPR reduction approach shown in Fig. 7 (b), with very small l.
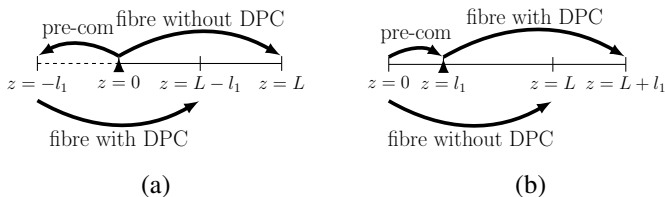

Fig. 7. Pre-equalization methods used (a) in [34] and (b) in Sec. IV.

C. Stochastic Memory

Consider the single-input multi-output channel $s_0^0 \mapsto \{\hat{s}_\ell^0\}_\ell$, where $\{\hat{s}_\ell^0\}_\ell$ denotes output symbols for the user-of-interest after equalization. Using the chain rule for mutual information, we find

$$I(s_0^0; \{\hat{s}_\ell^0\}_\ell) = I(s_0^0; \hat{s}_0^0) + \underbrace{I(s_0^0; \{\hat{s}_\ell^0\}_{\ell\neq0}|\hat{s}_0^0)}_{\text{stochastic memory}}. \qquad (7)$$

The stochastic memory refers to a component of the memory (intra-channel interactions) that is a function of noise; namely, it vanishes when noise is set to zero. This is the second term in (7) and is determined by signal-noise interactions. The stochastic memory is signal-dependent, growing with the input power, and cannot be equalized by digital signal processing at the receiver.

The stochastic memory causes the symbol energy to flow from s_0^0 to $\hat{s}_\ell^0$, $\ell \neq 0$. As $|s_0^0|$ is increased, the stochastic memory grows because the energy in $\{\hat{s}_\ell^0\}_{\ell\neq0}$ is increased, leading to a smaller $|\hat{s}_0^0|$; see the outer ring in Fig. 5 (b) and Fig. 6 (a). This amounts to a reduced SNR at the receiver for the channel $s_0^0 \mapsto \hat{s}_0^0$. Since only the first term in (7) is used for computing the AIR in this paper, a loss of SNR translates to a loss of AIR. Therefore, the stochastic memory, if not accounted for, incurs a penalty on the NFDM AIR at high powers.

The stochastic memory is investigated in a simulation with 5 users and (total linear) bandwidth of 60 GHz. We compare the extent of the stochastic memory in the fiber channel in the focusing regime with that in an AWGN channel whose SNR equals the SNR of the fiber channel. It is shown in Fig. 6 that the received symbols $|\hat{s}_0^0|$ in NFDM cluster in a cloud with a mean smaller than the transmit symbol $|s_0^0|$. This effect is the SNR (or energy) loss due to the stochastic memory. The signal-noise interaction in the focusing regime appears to be stronger than that in the defocusing regime.

The stochastic memory can be addressed by considering the single-input multiple-output channel $s_0^0 \mapsto \{\hat{s}_\ell^0\}_\ell$, instead of the single-input single-output channel $s_0^0 \mapsto \hat{s}_0^0$. Computing the second term in (7) will improve the NFDM AIR at high powers. It is expected that the AIR of the NFDM in Fig. 4 will not roll off at the maximum power $\mathcal{P} = 2.7$ dBm if maximum likelihood sequence detection is used at the Rx.

Note that the AIRs of different DoFs are not the same. Different users suffer from the stochastic memory to different extents. For example, the central user in WDM suffers more from nonlinear impairments.

V. IMPACT OF PERTURBATIONS AND NON-IDEALITIES

Research on NFDM has so far presented proof-of-concepts, demonstrating NFDM in ideal or simplified models. There remain many non-idealities and practical constraints, some of which are studied in this section. Specifically, we study the impacts of the loss, periodic amplification, third-order dispersion, and polarization-mode dispersion.

A. Loss with Periodic Amplification

It is often assumed that the fiber loss can be perfectly compensated using distributed Raman amplification. However, many optical systems are based on lumped amplification using erbium-doped fiber amplifiers (EDFAs). Furthermore, loss is not completely canceled by Raman amplification. There is typically some amplitude variation along the link. Consider an

optical fiber link with periodic lumped amplification. Signal propagation in one span can be described by

$$j\frac{\partial Q}{\partial z} = -j\frac{\alpha}{2}Q + \frac{\beta_2}{2}\frac{\partial^2 Q}{\partial t^2} + s\gamma|Q|^2Q, \qquad (8)$$

where α is the attenuation constant and γ is the nonlinearity coefficient; see Tab. I. Lumped amplification is performed at the end of each span of length $\mathcal{L}_{sp} = \mathcal{L}/N_{sp}$, where N_{sp} is the number of spans. After this, amplified spontaneous emission (ASE) noise is added to the signal, which is usually modeled as a white complex circular symmetric Gaussian stochastic process, with the well-known power spectral density (PSD) for EDFAs; see [35, Sec. 3. 1.3].

Introducing a change of variable

$$A(t,z) = e^{\frac{\alpha}{2}z}Q(t,z),$$

equation (8) is transformed to

$$j\frac{\partial A}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + s\gamma(z)|A|^2A, \qquad (9)$$

where $\gamma(z) = \gamma e^{-\alpha z}$. Approximating $\gamma(z)$ with its average $\bar{\gamma}$ in the interval $0 \le z \le \mathcal{L}_{sp}$ we obtain

$$j\frac{\partial A}{\partial z} \approx \frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + s\bar{\gamma}|A|^2A, \qquad (10)$$

where

$$\bar{\gamma} = \frac{\gamma}{\alpha\mathcal{L}_{sp}}(1 - e^{-\alpha\mathcal{L}_{sp}}).$$

Using parameters in Tab. I and II, we have $\bar{\gamma} = 0.3364$ $W^{-1}.km^{-1}$. Clearly, (10) can be normalized to (1) according to the change of variable in [2, above Eq. 3] with the scale factor $P = 2/(\bar{\gamma}\mathcal{L}) = 37.16 \times 10^{-4}$ W.

As a result, the non-integrable lossy model with periodic lumped amplification (8) is *approximated* by the integrable lossless path-averaged model (10). This approach, dating back to soliton communication, is successfully applied in [36] to address the loss problem in NFDM.

The error between the exact model (8) and the path-averaged model (10) grows linearly with distance and power [36]. It is not clear how an NFDM system designed using the averaged model (10) works in realistic EDFA-based systems represented by the original model (8). For example, although there is no interference in the ideal model in Fig. 3 (b), loss introduces interference as illustrated in Fig. 3 (d). To examine the AIRs, we simulate an optical link including loss and periodic amplification using the parameters found in Tab. II. The WDM and NFDM received constellations with 4000 noise realizations are displayed in Fig. 8. It is observed that NFDM is subject to a more phase distortion compared with NFDM in the ideal integrable model (compare Fig. 8(a) with Fig. 5 (b)). Nevertheless, the sizes of the noise clouds in NFDM are still smaller than those in WDM. This improvement then translates to a higher AIR for NFDM.

The AIRs of NFDM and WDM in the focusing regime with loss and periodic amplification are shown in Fig. 9. Although the AIR of NFDM also rolls off at high powers (potentially because the path-averaged model breaks down at high power), it reaches a higher maximum than the maximum WDM AIR.
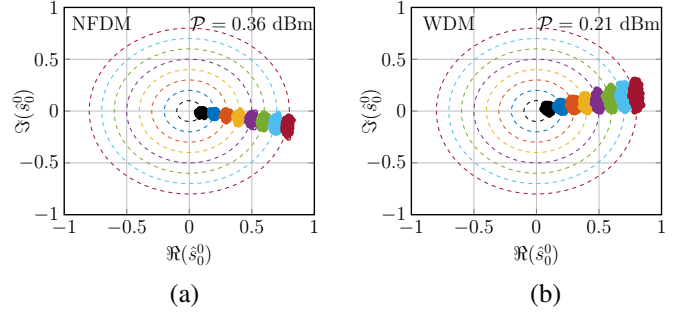

Fig. 8. Received constellations of (a) NFDM and (b) WDM systems in the focusing regime for the non-integrable model.

TABLE II
PARAMETERS OF THE PERIODIC AMPLIFICATION MODEL

$\mathcal{L}_{sp}$	80 km	span length
N_{sp}	20	number of spans
N_F	4.5 dB	noise figure of EDFAs
L	1600 km	total length

We conclude that NFDM offers a marginal peak-AIR gain of 0.41 bits/2D compared with WDM when considering an optical link with loss and lumped amplification.

B. Polarization Effects and Higher-order Dispersion

Two signals modulated in two polarizations of light may travel at different speeds along the fiber because of the randomly varying fiber birefringence. This leads to a temporal pulse broadening that is known as polarization-mode dispersion (PMD). There is also third-order dispersion, which can cause temporal broadening if the signal bandwidth is large. These two effects have not been accounted for so far.

The propagation of two signals u_x and u_y in the two polarizations of light can be described by the Manakov-PMD equation [37], [38, Eq.(12)]

$$\frac{\partial \vec{U}}{\partial \ell} + \frac{\alpha}{2}\vec{U} + \frac{j\beta_2}{2}\frac{\partial^2 \vec{U}}{\partial \tau^2} + \frac{\beta_3}{6}\frac{\partial^3 \vec{U}}{\partial \tau^3} - j\frac{8}{9}\gamma\vec{U}\|\vec{U}\|^2$$
$$= -\sigma_1\Delta\beta_1\frac{\partial \vec{U}}{\partial \tau}. \qquad (11)$$

Here, vector $\vec{U} = [u_x, u_y]^T$ contains signals in the x and y polarizations as a function of distance ℓ and time τ (in a certain frame co-propagating with signal [38]) and σ_1 is a Pauli matrix [37, Eq. (16), with $\alpha'(z) = 0$], given by

$$\sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Finally, β_2 and β_3 are the second- and third-order dispersion coefficients (that are the same for the x and y polarizations) and $\Delta\beta_1 = \frac{1}{2}(\beta_{1x} - \beta_{1y})$ is the differential group delay (DGD) parameter, where β_{1x} and β_{1y} are the first-order dispersion coefficients, respectively, for the x and y polarizations. In practice, $\Delta\beta_1$ varies randomly with distance (see below) and is responsible for the PMD.

The Manakov-PMD equation (11) is obtained from the coupled nonlinear Schrödinger equation, under a change of coordinate system [37]. To account for this in our simulations,
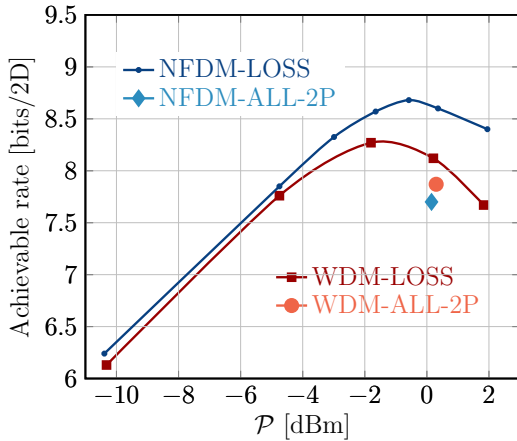
Fig. 9. AIRs of WDM and NFDM in non-integrable models. The curves denoted by LOSS correspond to the lossy model with periodic amplification (no PMD). The points denoted by ALL include all perturbations (loss, PMD and third-order dispersion).

TABLE III
POLARIZATION AND DISPERSION PARAMETERS

β_3	$0.06~\mathrm{ps}^3 \cdot \mathrm{m}^{-1}$	third-order dispersion coefficient
D_P	$0.1~\mathrm{ps}/\sqrt{\mathrm{km}}$	PMD parameter
ℓ_c	1 km	section length ℓ_c

we divide the fiber into a large number of small sections, each with length ℓ_c. Each small section is simulated according to (11). At the end of every section, we multiply $\vec{U}$ by a complex random unitary matrix

$$\mathcal{R} = e^{j\zeta} \begin{bmatrix} e^{j\psi} \cos\theta & e^{j\phi} \sin\theta \\ -e^{-j\phi} \sin\theta & e^{-j\psi} \cos\theta \end{bmatrix},$$

where ζ, ψ, ϕ are drawn independently and uniformly from $[0, 2\pi)$, and $\theta = \sin^{-1}(\sqrt{a})$, where a is drawn uniformly from $[0, 1]$.

The average amount of pulse broadening due to DGD is approximately $D_p\sqrt{\mathcal{L}}$, where $\mathcal{L}$ is the (total) fiber length and D_p is the PMD parameter [38], [39]. It is customary to draw the random variable $\Delta\beta_1$ at the end of each section from a real Gaussian distribution with mean $\mu_p = D_p\sqrt{\ell_c}$ and variance $\sigma_P = 0.2\mu_p$ [40]. $\Delta\beta_1$ changes independently from section to section.

The impacts of the PMD and higher-order dispersion on WDM are well studied. It is shown in [40] that the performance improvement in applying BP to WDM saturates due to PMD as the number of back-propagating channels is increased. On the other hand, the impact of these effects on NFDM has not been investigated yet. To clarify this, we carry out a simulation, with parameters in Tab. III.

We assume that the receiver has the channel state information. Under this assumption, all random rotations and random DGDs during propagation are compensated at the receiver in both schemes. Note that while the effects of β_2 and γ are equalized in the nonlinear Fourier domain via the channel filter (5), the third-order dispersion and nonlinear interactions between the two polarizations act as distortions in NFDM.
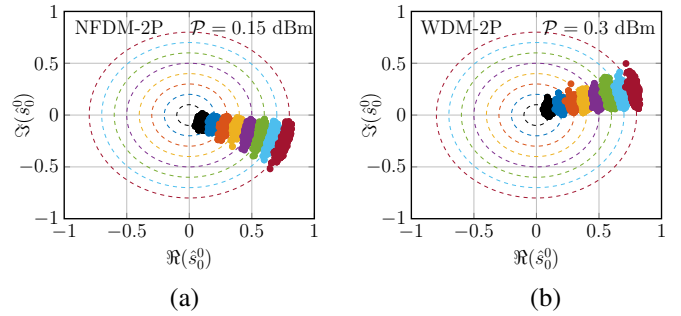

Fig. 10. Received constellations of NFDM and WDM in the focusing regime with PMD and third-order dispersion, in dual polarization transmission.

In WDM, the UOI is filtered in the frequency domain in both u_x and u_y and the resulting vector is back-propagated according to (11). *This includes the inter-polarization nonlinearity mitigation (jointly across x and y) and the third-order dispersion compensation.* Fig. 10 shows the received constellations for WDM and NFDM with two polarizations. The corresponding AIRs are shown in Fig. 9. There is a performance degradation in both multiplexing schemes compared to the single polarization case. The AIR of the WDM and NFDM are nearly equal when all perturbations are included. However, full dual-polarization BP was applied at the Rx for WDM, which includes equalizing the in-band inter-polarization nonlinear effects, the third-order dispersion and the per-span loss. These perturbations were not compensated in NFDM, because the channel filter $\exp(4js\lambda^2 z)$ in (5) includes only the second-order dispersion and nonlinearity. Compensating the above perturbations in the time domain is expected to improve NFDM.

We conclude that the AIR of the independent single-polarization NFDM transmission and detection (not using joint NFT) subject to third-order dispersion and loss is nearly equal to the WDM AIR with joint dual-polarization back-propagation and compensation of the third-order dispersion and loss, in a system with $N_u = 5$ and $N_s = 1$. We anticipate that the NFDM AIR will be higher than the WDM AIR in comparable systems with joint dual-polarization transmission and detection and compensation of the same perturbations. In fact, it has been shown recently that if joint (de)-modulation is performed across two polarizations in NFDM using the NFT of the two-dimensional signals, the NFDM outperforms WDM in terms of Q-factor [41].

The algorithm used in this paper is a simple discrete layer-peeling method. More accurate algorithms exist that allow operation at higher powers and reduce the contribution of the numerical error to the AIR, which is appreciable at high powers. These better algorithms, such as that in [42], likely improve NFDM and its AIR presented in this paper.

C. Spectral Efficiency

In the simulations thus far, we considered one symbol per user and calculated the data rates in bits/DoF. In this section, we compute the spectral efficiency (SE) in bits/s/Hz.

Numerical computation of the SE is difficult because the number of symbols N_s should be sufficiently large so that the
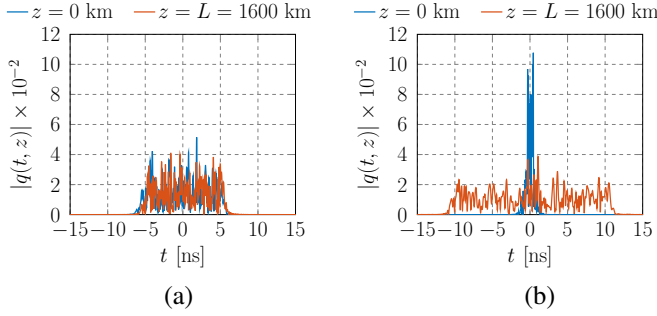
Fig. 11.   NFDM signals (a) with and (b) without pre-compensation.

ratio r_g of the guard-time to the blocklength is small. With $N_s = 1$, the average power and the SE are low in both WDM and NFDM. Consequently, in the literature the SE is usually computed by considering the main lobe of the pulse shape (eg., a root raised-cosine function) as the symbol duration. The impact of the pulse broadening in time due to dispersion on the SE is typically ignored because it is assumed that $r_g \to 0$ if $N_s \to \infty$. The WDM SEs in the literature are per-sample mutual information as in Fig. 4, without taking into account guard-times.

The data rates shown in Fig. 4 were obtained for the same time duration and bandwidth in WDM and NFDM. As a result, the gain in SE in bits/s/Hz (ratio of the SE of the NFDM and WDM) is same as the gain in AIR in bits/DoF. Nevertheless, its is still instructive to see the absolute numbers for the approximate SE as well. To do so we consider a system with $N_u = 5$ users, $N_s = 15$ and parameters in Tab. II. The pre-compensation technique in [34] is used to reduce the guard time; see Fig. 11 (a)–(b).

Fig. 12 shows the received symbols in the NFDM and WDM system. It can be seen that the size of clouds is bigger for $N_s = 15$ compared to $N_s = 1$. This effect is attributed to the growth of the signal-noise interaction with the number of DoFs, which is fundamental and holds for both WDM and NFDM [32, Theorem 1]. The SEs of the NFDM and WDM system are shown in Fig. 13. The split-BP is applied in WDM simulations so that the signals in the two systems have roughly the same temporal broadening. The WDM SE is a little greater than NFDM SE. We attribute this to the fact that the NFT algorithm, as well as potentially the path-averaged loss model, become less accurate as N_s is increased.

For clarity, we also compare time duration T, bandwidth $\mathcal{W}$ and guard-time T_g requirement of NFDM and WDM in Fig. 14. It can be seen that, for a fixed T, $\mathcal{W}$ and $\mathcal{P}$ at the input, T and $\mathcal{W}$ at the output, and thus T_g, are approximately the same in both schemes. The maximum input output time duration and bandwidth is about the same for all values of powers, around 11 ns and 96 GHz.

A SE of 2 bits/s/Hz was reported in [43], however the distance and bandwidth are half of their values in this paper. Doubling the distance will double the guard-time, and also introduce more noise. Although theoretically linear and nonlinear multiplexing should require approximately the same guard-time, in practice r_g is lower for WDM because transmission is possible for larger N_s. This is however due to the
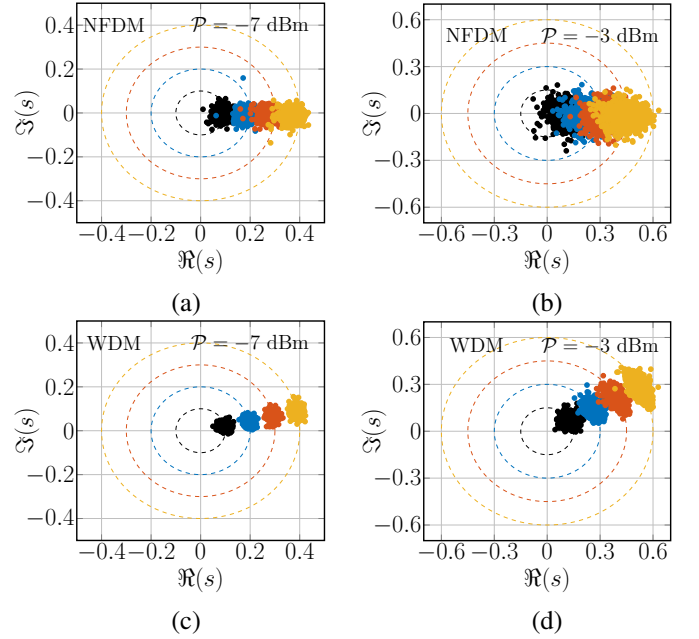

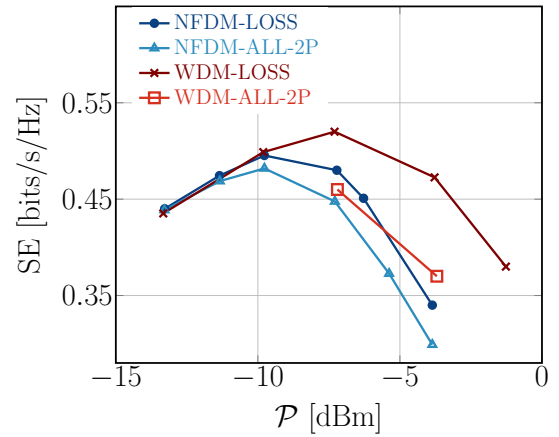Fig. 12.   Received symbols in NFDM and WDM with $N_s = 15$.

Fig. 13.   The SE of NFDM and WDM with $N_s = 15$ and guard-time. For dual-polarization transmissions in both systems, the SE and power is measured per polarization.

lower complexity advantage of WDM, and not a fundamental performance difference.

The SE may be maximized by generating compact signals using the periodic NFT [22] or the standard NFT with a modified modulation [44].

Remark 1. We caution that the SE in the short blocklength regime depends on the signal set. For a fixed average power and time duration T, different signal sets can have very different SEs. We have not maximized SEs at each power in Fig. 13. Thus, it is difficult to draw definite conclusions on SEs based on Fig. 13, because the SE appeared quite sensitive in our simulations.

VI.   COMMENTS ON COMPLEXITY

The computational complexity of the NFT with continuous spectrum in terms of the number of samples N in time is $O[N \log^2 N]$ [45]. Lima et al. investigated the required number
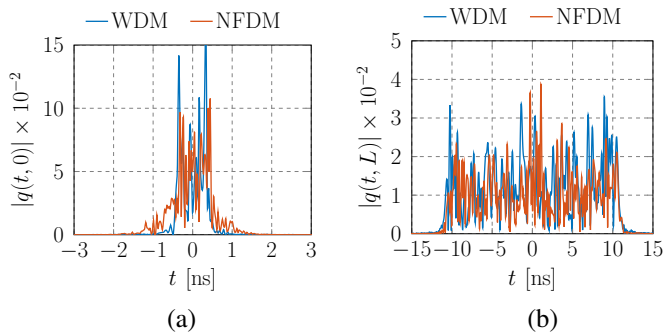
Fig. 14. Comparison of the time duration at the (a) input and (b) output. Here $N_u = 5$ and $N_s = 15$. The time duration at input is $T_{in} \approx 2.1\ ns$, while $T_{out} \approx 22.3\ ns$, which can be reduced to 11 ns as shown in Fig. 11. The temporal expansion ratio for $N_s = 15$ is 10.

TABLE IV
COMPLEXITY OF NFDM AND WDM

	$\mathcal{P}$[dBm]	N_s	N	E_s	FLOP
NFDM	0.33	2^7	2^{17}	1.13%	2.23×10^{11}
WDM	0.33	2^7	2^8	0.24%	6.93×10^6
NFDM	-0.9	2^6	2^{14}	1.16%	3.49×10^9
WDM	-0.5	2^6	2^7	0.24%	3.46×10^6

TABLE V
MAXIMUM POWER REACHED BY THE BASIC DLP

$\mathcal{P}$[dBm]	N_s	N	E_s
0.33	2^7	2^{17}	1.13%
-0.9	2^6	2^{14}	1%

of samples in the nonlinear frequency per processing frame (resolution spectrum) in terms of the average power in the defocusing regime [42]. The processing frame is defined as the pulse duration at the Rx, which is typically longer than pulse duration at the Tx. It is argued that a prohibitively fine spectral resolution is needed at high powers for NFDM to be viable in practice. The algorithm used in [42] is the discrete layer-peeling (DLP) method used in this paper, but together with applying the GLM integral equations in each small step in time. It is known that such refinements of the DLP reduce the numerical error, albeit with increased complexity.

To reduce complexity, it is desirable to minimize pulse broadening and the inter-frame guard times. We apply the pre-equalization proposed by Tavakkolnia et al. (with $\ell_1 = \mathcal{L}/2$). This causes the signal to contract in the time domain as it evolves in $0 \leq z \leq \mathcal{L}/2$, and subsequently broaden as it evolves in $\mathcal{L}/2 \leq z \leq \mathcal{L}$. Consequently, the signal durations at the Tx and Rx are similar.

Tab. IV compares the number of samples N in the nonlinear frequency domain and N_s in NFDM and WDM for one user with an oversampling rate of 3; see (4). The NFT and INFT are implemented by the basic forward and inverse DLP algorithms in the focusing regime, whose details can be found in [32]. The complexity of the BP algorithm using the split-step Fourier method is evidently $O[MN \log N]$, where M is the number of segments in distance. In our simulation $M = 1600$, corresponding to a step size of 1 km. The complexity of the two algorithms is compared in terms of the total number of floating-point operations (FLOPs), which may depend on details of the implementations. The error E_s is the symbol error defined as $\|\{s_\ell^k\} - \{\hat{s}_\ell^k\}\|/\|\{s_\ell^k\}\|$, where $\{s_\ell^k\}$ and $\{\hat{s}_\ell^k\}$ are defined in Sec. III.

The NFT algorithms typically work up to a certain power before numerical error becomes significant. Tab. V shows the maximum power that can be reached using the basic forward and inverse DLP with pre-equalization.

It can be concluded that the complexity of the NFT with the DLP algorithm is substantially higher than the complexity of the FFT, and rapidly grows with N_s. This seems to be consistent with the findings of [42] too. The reader is referred to [42], in which the complexity and the maximum power are studied in much more detail.

Note that NFDM currently requires successive NFT and INFT operations at ADMs digitally in the electrical domain (thus requiring optical-to-electrical conversion). In WDM, these operations can be done easily using optical signal processing (without expensive optical-to-electrical conversion). As a result, NFDM is currently impractical. Implementing NFDM using optical signal processing is an interesting and important research direction.

VII. CONCLUSIONS

We compared the AIRs of the WDM and NFDM. It is shown that: (i) the NFDM AIR is greater than the WDM AIR in an ideal integrable model with five users and one symbol per user; (ii) the AIR of the independent single-polarization NFDM transmission and detection (not using joint NFT) subject to third-order dispersion and per-span loss is approximately equal to the WDM AIR with joint dual-polarization back-propagation together with the compensation of the third-order dispersion and per-span loss, in a system with five users and one symbol per user. We anticipate that the NFDM AIR will be higher than the WDM AIR in this particular system if joint dual-polarization transmission and detection, and compensation of the same perturbations, in applied in NFDM.

ACKNOWLEDGMENTS

REFERENCES

[1] A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Eur. Conf. Opt. Commun.*, Montreux, Switzerland, Sep. 1993, pp. 41–44.

[2] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Part I: Mathematical tools," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4312–4328, Jul. 2014.

[3] ——, "Information transmission using the nonlinear Fourier transform, Part II: Numerical methods," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4329–4345, Jul. 2014.

[4] ——, "Information transmission using the nonlinear Fourier transform, Part III: Spectrum modulation," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4346–4369, Jul. 2014.

[5] J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, "Nonlinear spectral management: Linearization of the lossless fiber channel," *Opt. Exp.*, vol. 21, no. 20, pp. 24 344–24 367, Oct. 2013.

[6] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," *Opt. Exp.*, vol. 22, no. 22, pp. 26 720–26 741, Nov. 2014.

[7] S. Wahls, S. T. Le, J. E. Prilepsk, H. V. Poor, and S. K. Turitsyn, "Digital backpropagation in the nonlinear Fourier domain," in *16th Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Stockholm, Sweden, Jun. 2015, pp. 445–449.

[8] V. Aref, S. T. Le, and H. Buelow, "Demonstration of fully nonlinear spectrum modulated system in the highly nonlinear optical transmission regime," in *Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, Sep. 2016, pp. 1–3.

[9] I. Tavakkolnia and M. Safari, "Signalling over nonlinear fibre-optic channels by utilizing both solitonic and radiative spectra," in *Eur. Conf. Networks and Commun.*, Paris, France, Jul. 2015, pp. 103–107.

[10] S. T. Le *et al.*, "Demonstration of nonlinear inverse synthesis transmission over transoceanic distances," *J. Lightw. Technol.*, vol. 34, no. 10, pp. 2459–2466, May 2016.

[11] S. T. Le, V. Aref, and H. Bülow, "Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit," *Nature Photon.*, vol. 11, pp. 570–576, Jul. 2017.

[12] S. Hari, M. I. Yousefi, and F. R. Kschischang, "Multieigenvalue communication," *J. Lightw. Technol.*, vol. 34, no. 13, pp. 3110–3117, Jul. 2016.

[13] Z. Dong *et al.*, "Nonlinear frequency division multiplexed transmissions based on NFT," *IEEE Photon. Technol. Lett.*, vol. 27, no. 15, pp. 1621–1623, Aug. 2015.

[14] T. Gui *et al.*, "4 bits/symbol phase and amplitude modulation on a single discrete eigenvalue for transmissions based on nonlinear Fourier transform," in *Opt. Fiber Commun. Conf. and Exhib.*, Los Angeles, CA, USA, Mar. 2017, pp. 1–3.

[15] S. T. Le, H. Buelow, and V. Aref, "Demonstration of 64× 0.5Gbaud nonlinear frequency division multiplexed transmission with 32QAM," in *Opt. Fiber Commun. Conf. and Exhib.*, Los Angeles, CA, USA, Mar. 2017, pp. 1–3.

[16] A. Geisler and C. G. Schaeffer, "Experimental nonlinear frequency division multiplexed transmission using eigenvalues with symmetric real part," in *Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, Sep. 2016, pp. 1–3.

[17] H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," in *Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, Sep. 2016, pp. 1–3.

[18] I. Tavakkolnia and M. Safari, "Capacity analysis of signaling on the continuous spectrum of nonlinear optical fibers," *J. Lightw. Technol.*, vol. 35, no. 11, pp. 2086–2097, Jun. 2017.

[19] S. K. Turitsyn *et al.*, "Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives," *Optica*, vol. 4, no. 3, pp. 307–322, Mar. 2017.

[20] T. Gui, T. H. Chan, C. Lu, A. P. T. Lau, and P. K. A. Wai, "Alternative decoding methods for optical communications based on nonlinear Fourier transform," *J. Lightw. Technol.*, vol. 35, no. 9, pp. 1542–1550, May 2017.

[21] S. T. Le *et al.*, "Modified nonlinear inverse synthesis for optical links with distributed raman amplification," in *Eur. Conf. Opt. Commun.*, Valencia, Spain, Sep. 2015, pp. 1–3.

[22] M. K. Kopae, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn, "Optical communication based on the periodic nonlinear Fourier transform signal processing," in *IEEE 6th Int. Conf. Photon. (ICP)*, Kuching, Malaysia, Mar. 2016, pp. 1–3.

[23] M. I. Yousefi and X. Yangzhang, "Linear and nonlinear frequency-division multiplexing," in *Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, Sep. 2016, pp. 1–3.

[24] ——, "Linear and nonlinear frequency-division multiplexing," arXiv:1603.04389v2, May 2016.

[25] X. Yangzhang, M. I., A. Alvarado, D. Lavery, and P. Bayvel, "Nonlinear frequency-division multiplexing in the focusing regime," in *Opt. Fiber Commun. Conf. and Exhib.*, Los Angeles, CA, USA, Mar. 2017, pp. 1–3.

[26] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fiber communications," *Lett. Nature*, vol. 411, pp. 1027–1030, Jun. 2001.

[27] J. Tang, "The channel capacity of a multispan DWDM system employing dispersive nonlinear optical fibers and an ideal coherent optical receiver," *J. Lightw. Technol.*, vol. 20, no. 7, pp. 1095–1101, Nov. 2002.

[28] I. B. Djordjevic, B. Vasic, M. Ivkovic, and I. Gabitov, "Achievable information rates for high-speed long-haul optical transmission," *J. Lightw. Technol.*, vol. 23, no. 11, p. 3755, Dec. 2005.

[29] R. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[30] M. I. Yousefi, "The Kolmogorov-Zakharov model for optical fiber communication," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 377 – 391, Jan. 2017.

[31] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, "Capacity estimates for optical transmission based on the nonlinear Fourier transform," *Nature Commun.*, vol. 7, no. 12710, Sep. 2016.

[32] M. I. Yousefi, "The asymptotic capacity of the optical fiber," arXiv:1610.06458, pp. 1–12, Oct. 2016.

[33] D. Lavery *et al.*, "The benefit of split nonlinearity compensation for single-channel optical fiber communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 17, pp. 1803–1806, Sep. 2016.

[34] I. Tavakkolnia and M. Safari, "Dispersion pre-compensation for NFT-based optical fiber communication systems," in *Conf. Laser and Electro-Optics*, San Jose, CA, USA, Jun. 2016, pp. 4–5.

[35] L. F. Mollenauer and J. P. Gordon, *Solitons in Optical Fibers: Fundamentals and Applications*, 1st ed. Amsterdam, The Netherlands: Elsevier Academic Press, 2006.

[36] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," *Opt. Exp.*, vol. 23, no. 7, pp. 8317–8328, Mar. 2015.

[37] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1745, Sep. 1997.

[38] G. Agrawal, *Nonlinear Fiber Optics*, 4th ed. San Francisco, CA, USA: Academic Press, 2007.

[39] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 14, no. 2, pp. 148–157, Feb. 1996.

[40] G. Liga *et al.*, "Ultra-wideband nonlinearity compensation performance in the presence of PMD," in *Eur. Conf. Opt. Commun.*, Düsseldorf, Germany, Sep. 2016, pp. 1–3.

[41] J.-W. Goossens, M. I. Yousefi, Y. Jaouën, and H. Hafermann, "Polarization-division multiplexing based on the nonlinear Fourier transform," arXiv:1707.08589, Jul. 2017.

[42] I. T. Lima, V. S. Grigoryan, M. O'Sullivan, and C. R. Menyuk, "Computational complexity of nonlinear transforms applied to optical communications systems with normal dispersion fibers," in *IEEE Photo. Conf., IPC 2015*, vol. 3, Reston, VA, USA, Oct. 2015, pp. 277–278.

[43] S. T. Le, V. Aref, and H. Buelow, "125 Gbps pre-compensated nonlinear frequency-division multiplexed transmission," in *Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, pp. 1–3.

[44] S. Wahls, "Generation of time-limited signals in the nonlinear Fourier domain via b-modulation," in *Eur. Conf. Opt. Commun.*, Gothenburg, Sweden, Sep. 2017, pp. 1–3.

[45] V. Vaibhav, "Fast inverse nonlinear Fourier transformation using exponential one-step methods, Part I: Darboux transformation," arXiv:1704.00951, Apr. 2017.